Introduction to Econometrics
(3rd Updated Edition, Global Edition)

by

James H. Stock and Mark W. Watson

Solutions to End-of-Chapter Exercises: Chapter 2*

(This version August 17, 2014)

*Limited distribution: **For Instructors Only**. Answers to all odd-numbered questions are provided to students on the textbook website. If you find errors in the solutions, please pass them along to us at mwatson@princeton.edu.

2.1. (a) Probability distribution function for Y

Outcome (number of heads)	$Y = 0$	$Y = 1$	$Y = 2$
Probability	0.36	0.48	0.16

(b) $\mu_Y = E(Y) = (0\times 0.36) + (1\times 0.48) + (2\times 0.16) = 0.8.$ $F \xrightarrow{d} Fq, \infty.$

Using Key Concept 2.3: $\text{var}(Y) = E(Y^2) - [E(Y)]^2,$

and

$$E(Y^2) = \sigma_Y{}^2 + \mu_Y{}^2 = 2\times 0.6\times 0.4 + 0.8^2 = 1.12$$

so that

$$\text{var}(Y) = E(Y^2) - [E(Y)]^2 = 1.12 - 0.8^2 = 0.48.$$

2.2. We know from Table 2.2 that $\Pr(Y=0)=0.22$, $\Pr(Y=1)=0.78$, $\Pr(X=0)=0.30$, $\Pr(X=1)=0.70$. So

(a)

$$\begin{aligned}\mu_Y &= E(Y) = 0\times\Pr(Y=0)+1\times\Pr(Y=1)\\ &= 0\times 0.22+1\times 0.78 = 0.78,\end{aligned}$$

$$\begin{aligned}\mu_X &= E(X) = 0\times\Pr(X=0)+1\times\Pr(X=1)\\ &= 0\times 0.30+1\times 0.70 = 0.70.\end{aligned}$$

(b)

$$\begin{aligned}\sigma_X^2 &= E[(X-\mu_X)^2]\\ &= (0-0.70)^2\times\Pr(X=0)+(1-0.70)^2\times\Pr(X=1)\\ &= (-0.70)^2\times 0.30+0.30^2\times 0.70 = 0.21,\\ \sigma_Y^2 &= E[(Y-\mu_Y)^2]\\ &= (0-0.78)^2\times\Pr(Y=0)+(1-0.78)^2\times\Pr(Y=1)\\ &= (-0.78)^2\times 0.22+0.22^2\times 0.78 = 0.1716.\end{aligned}$$

(c)

$$\begin{aligned}\sigma_{XY} &= \operatorname{cov}(X,Y) = E[(X-\mu_X)(Y-\mu_Y)]\\ &= (0-0.70)(0-0.78)\Pr(X=0, Y=0)\\ &\quad +(0-0.70)(1-0.78)\Pr(X=0, Y=1)\\ &\quad +(1-0.70)(0-0.78)\Pr(X=1, Y=0)\\ &\quad +(1-0.70)(1-0.78)\Pr(X=1, Y=1)\\ &= (-0.70)\times(-0.78)\times 0.15+(-0.70)\times 0.22\times 0.15\\ &\quad +0.30\times(-0.78)\times 0.07+0.30\times 0.22\times 0.63\\ &= 0.084,\end{aligned}$$

$$\operatorname{corr}(X,Y) = \frac{\sigma_{XY}}{\sigma_X\sigma_Y} = \frac{0.084}{\sqrt{0.21\times 0.1716}} = 0.4425.$$

2.3. For the two new random variables $W = 4+8X$ and $V = 11-2Y$, we have:

(a)

$$E(V) = E(11-2Y) = 11-2E(Y) = 11-2\times 0.78 = 9.44,$$
$$E(W) = E(4+8X) = 4+8E(X) = 4+8\times 0.70 = 9.6.$$

(b)

$$\sigma_W^2 = \text{var}\,(4+8X) = 8^2\sigma_X^2 = 64\times 0.21 = 13.44,$$
$$\sigma_V^2 = \text{var}\,(11+2Y) = (-2)^2\sigma_Y^2 = 4\times 0.1716 = 0.6864.$$

(c)

$$\sigma_{WV} = \text{cov}\,(4+8X, 11+2Y) = 6\times(-2)\,\text{cov}\,(X, Y) = -12\times 0.084 = -1.008$$

$$\text{corr}\,(W,\ V) = \frac{\sigma_{WV}}{\sigma_W\sigma_V} = \frac{-1.008}{\sqrt{13.44\times 0.6864}} = -0.3319$$

2.4. (a) $E(X^3) = 0^3 \times (1-p) + 1^3 \times p = p$

(b) $E(X^k) = 0^k \times (1-p) + 1^k \times p = p$

(c) $E(X) = 0.3$

$\text{var}(X) = E(X^2) - [E(X)]^2 = 0.3 - 0.09 = 0.21$

Thus, $\sigma = \sqrt{0.21} = 0.46.$

To compute the skewness, use the formula from exercise 2.21:

$$\begin{aligned} E(X-\mu)^3 &= E(X^3) - 3[E(X^2)][E(X)] + 2[E(X)]^3 \\ &= 0.3 - 3\times 0.3^2 + 2\times 0.3^3 = 0.084 \end{aligned}$$

Alternatively, $E(X-\mu)^3 = [(1-0.3)^3 \times 0.3] + [(0-0.3)^3 \times 0.7] = 0.084$

Thus, skewness $= E(X-\mu)^3/\sigma^3 = .084/0.46^3 = 0.87.$

To compute the kurtosis, use the formula from exercise 2.21:

$$\begin{aligned} E(X-\mu)^4 &= E(X^4) - 4[E(X)][E(X^3)] + 6[E(X)]^2[E(X^2)] - 3[E(X)]^4 \\ &= 0.3 - 4\times 0.3^2 + 6\times 0.3^3 - 3\times 0.3^4 = 0.0777 \end{aligned}$$

Alternatively, $E(X-\mu)^4 = [(1-0.3)^4 \times 0.3] + [(0-0.3)^4 \times 0.7] = 0.0777$

Thus, kurtosis is $E(X-\mu)^4/\sigma^4 = .0777/0.46^4 = 1.76$

2.5. Let X denote temperature in °F and Y denote temperature in °C. Recall that $Y = 0$ when $X = 32$ and $Y = 100$ when $X = 212$.

This implies $Y = (100/180) \times (X - 32)$ or $Y = -17.78 + (5/9) \times X$.

Using Key Concept 2.3, $\mu_X = 65^{o}$F implies that $\mu_x = -17.78 + (5/9) \times 63 = 18.33$°C, and $\sigma_X = 5^{o}$F implies $\sigma_x = (5/9) \times 5 = 2.78$°C.

2.6. The table shows that $\Pr(X=0, Y=0) = 0.078$, $\Pr(X=0, Y=1) = 0.673$, $\Pr(X=1, Y=0) = 0.042$, $\Pr(X=1, Y=1) = 0.207$, $\Pr(X=0) = 0.751$, $\Pr(X=1) = 0.249$, $\Pr(Y=0) = 0.12$, $\Pr(Y=1) = 0.88$.

(a)

$$\begin{aligned} E(Y) = \mu_Y &= 0 \times \Pr(Y=0) + 1 \times \Pr(Y=1) \\ &= 0 \times 0.12 + 1 \times 0.88 = 0.88 \end{aligned}$$

(b)

$$\text{Unemployment Rate} = \frac{\#(\text{unemployed})}{\#(\text{labor force})} = \Pr(Y=0) = 0.12$$

(c) Calculate the conditional probabilities first:

$$\Pr(Y=0|X=0) = \frac{\Pr(X=0, Y=0)}{\Pr(X=0)} = \frac{0.078}{0.751} = 0.104,$$

$$\Pr(Y=1|X=0) = \frac{\Pr(X=0, Y=1)}{\Pr(X=0)} = \frac{0.673}{0.751} = 0.896,$$

$$\Pr(Y=0|X=1) = \frac{\Pr(X=1, Y=0)}{\Pr(X=1)} = \frac{0.042}{0.249} = 0.169,$$

$$\Pr(Y=1|X=1) = \frac{\Pr(X=1, Y=1)}{\Pr(X=1)} = \frac{0.207}{0.249} = 0.831.$$

The conditional expectations are

$$\begin{aligned} E(Y|X=1) &= 0 \times \Pr(Y=0|X=1) + 1 \times \Pr(Y=1|X=1) \\ &= 0 + 1 \times 0.831 = 0.831 \end{aligned}$$

$$\begin{aligned} E(Y|X=0) &= 0 \times \Pr(Y=0|X=0) + 1 \times \Pr(Y=1|X=0) \\ &= 0 + 1 \times 0.896 = 0.896 \end{aligned}$$

(d) Use the solution to part (c),

Unemployment rate for college graduates = $1 - E(Y \mid X = 1) = 1 - 0.831 = 0.169$

Unemployment rate for non-college graduates = $1 - E(Y \mid X = 0) = 1 - 0.896 = 0.104$

(e) The probability that a randomly selected worker, who is reported being unemployed, is a college graduate is

$$\Pr(X = 1 \mid Y = 0) = \frac{\Pr(X = 1, Y = 0)}{\Pr(Y = 0)} = \frac{0.042}{0.12} = 0.35$$

The probability that this worker is a non-college graduate is

$$\Pr(X = 0 \mid Y = 0) = 1 - \Pr(X = 1 \mid Y = 0) = 1 - 0.35 = 0.65$$

(f) Educational achievement and employment status are not independent because they do not satisfy that, for all values of x and y,

$$\Pr(X = x \mid Y = y) = \Pr(X = x)$$

For example, from part (e) $\Pr(X = 0 \mid Y = 0) = 0.65$, while from the table $\Pr(X = 0) = 0.751$.

2.7. Using obvious notation, $C = M + F$; thus $\mu_C = \mu_M + \mu_F$ and $\sigma_C^2 = \sigma_M^2 + \sigma_F^2 + 2\,\text{cov}(M,\, F)$. This implies

(a) $\mu_C = 50 + 48 = \$98,000$ per year.

(b) $\text{corr}(M,\, F) = \frac{\text{cov}(M,\, F)}{\sigma_M \sigma_F}$, so that $\text{cov}(M,\, F) = \sigma_M \sigma_F \text{corr}(M,\, F)$. Thus, $\text{cov}(M,\, F) = 15 \times 13 \times 0.9 = 175.50,$ where the units are squared thousands of dollars per year.

(c) $\sigma_C^2 = \sigma_M^2 + \sigma_F^2 + 2\,\text{cov}(M, F),$ so that $\sigma_C^2 = 15^2 + 13^2 + 2 \times 175.5 = 745,$ and $\sigma_C = \sqrt{745} = 27.295$ thousand dollars per year.

(d) First you need to look up the current Euro/dollar exchange rate in the Wall Street Journal, the Federal Reserve web page, or other financial data outlet. Suppose that this exchange rate is e (say e = 0.75 Euros per Dollar); each 1 dollar is therefore with e Euros. The mean is therefore $e \times \mu_C$ (in units of thousands of Euros per year), and the standard deviation is $e \times \sigma_C$ (in units of thousands of Euros per year). The correlation is unit-free, and is unchanged.

2.8. $E(Y) = 4$; $\text{var}(Y) = 1/9$; $Z = 3(Y - 4)$

$$\mu_Z = E(3(Y-4)) = 3(E(Y)-4) = 3(4-4) = 0$$

$$\sigma_Z^2 = 9\text{var}(Y) = 9 \times \frac{1}{9} = 1$$

2.9.

		Value of *Y*					Probability Distribution of *X*
		2	**4**	**6**	**8**	**10**	
Value of *X*	**3**	0.04	0.09	0.03	0.12	0.01	0.29
	6	0.10	0.06	0.15	0.03	0.02	0.36
	9	0.13	0.11	0.04	0.06	0.01	0.35
Probability Distribution of *Y*		0.27	0.26	0.22	0.21	0.04	1.00

(a) The probability distribution is given in the table above.

$$E(Y) = 2(0.27) + 4(0.26) + 6(0.22) + 8(0.21) + 10(0.04) = 4.98$$

$$E(Y^2) = 2^2(0.27) + 4^2(0.26) + 6^2(0.22) + 8^2(0.21) + 10^2(0.04) = 30.6$$

$$\text{var}(Y) = E(Y^2) - [E(Y)]^2 = 30.6 - 24.8 = 5.8$$

$$\sigma_Y = \sqrt{5.8} = 2.41$$

(b) The conditional probability of $Y|X = 6$ is given in the table below

Value of *Y*				
2	**4**	**6**	**8**	**10**
0.10/0.36	0.06/0.36	0.15/0.36	0.03/0.36	0.02/0.36

$$E(Y|X = 6) = 2\times(0.1/0.36) + 4\times(0.06/0.36) + 6\times(0.15/0.36) + 8\times(0.03/0.36) + 10\times(0.02/0.36) = 4.944$$

$$E(Y^2|X = 6) = 2^2\times(0.1/0.36) + 4^2\times(0.06/0.36) + 6^2\times(0.15/0.36) + 8^2\times(0.03/0.36) + 10^2\times(0.02/0.36) = 29.667$$

$$\text{var}(Y) = E(Y^2) - [E(Y)]^2 = 29.667 - 24.443 = 5.244$$

(c)

$$E(XY) = (3\times2\times0.04) + (3\times4\times0.09) + \cdots(9\times10\times0.01) = 29.4$$

$$\text{cov}(X,Y) = E(XY) - E(X)E(Y) = 29.4 - 6.18\times4.98 = -1.3764$$

$$\text{corr}(X,Y) = \text{cov}(X,Y)/(\sigma_X\sigma_Y) = -1.3764/(2.93\times2.41) = -0.1949$$

2.10. Using the fact that if $Y \sim N\left(\mu_Y, \sigma_Y^2\right)$ then $\frac{Y-\mu_Y}{\sigma_Y} \sim N(0,1)$ and Appendix Table 1, we have

(a) $\Pr(Y \leq 5) = \Pr\left(\frac{Y-4}{3} \leq \frac{5-4}{3}\right) = \Phi\left(\frac{1}{3}\right) = 0.6304$

(b) $\Pr(Y > 2) = 1 - \Pr\left(\frac{Y-5}{4} \leq \frac{2-5}{4}\right) = 1 - \Phi\left(-\frac{1}{4}\right) = \Phi\left(\frac{1}{4}\right) = 0.5987$

(c) $\Pr(2 \leq Y \leq 5) = \Pr\left(\frac{2-1}{2} \leq \frac{Y-1}{2} \leq \frac{5-1}{2}\right) = \Phi(2) - \Phi\left(\frac{1}{2}\right) = 0.2857$

(d) $\Pr(1 \leq Y \leq 4) = \Pr\left(\frac{1-2}{1} \leq \frac{Y-2}{1} \leq \frac{1-2}{1}\right) = \Phi(2) - \Phi(-1)$

$$= \Phi(2) - (1 - \Phi(1) = 0.8185$$

2.11. (a) 0.90

(b) 0.95

(c) 0.95

(d) When $Y \sim \chi_8^2$, then $\frac{Y}{10} \sim F_{8,\infty}$

(e) $Y = Z^2$, where $Z \sim N(0,1)$, thus, $\Pr(Y \le 0.5) = \Pr(-0.5 \le Z \le 0.5) = 0.383$

2.12. (a) 0.90

(b) 0.90

(c) 0.9108

(d) The t_{df} distribution and $N(0, 1)$ approach each other only when N becomes large, or the degrees of freedom become large.

(e) 0.95

(f) 0.01

2.13. (a) $E(Y^2) = \text{Var}(Y) + \mu_Y^2 = 4 + 0 = 4$

$E(W^2) = \text{Var}(W) + \mu_W^2 = 16 + 0 = 16$

(b) Y and W are symmetric around 0, thus skewness is equal to 0; because their mean is 0, which means that the third moment is 0.

(c) The kurtosis of the normal is 3, so $3 = \frac{E(Y-\mu_Y)^4}{\sigma_Y^4}$; by transforming both Y and W to the standard normal yields the results.

(d) First, condition on $X = 0$, so that $S = W$:

$E(S|X=0) = 0; E(S^2|X=0) = 16, E(S^3|X=0) = 0, E(S^4|X=0) = 3\times 16^2$

Similarly,

$E(S|X=1) = 0; E(S^2|X=1) = 4, E(S^3|X=1) = 0, E(S^4|X=1) = 3\times 4^2$

From the law of iterated expectations

$$E(S) = E(S|X=0)\times \Pr(\text{X}=0) + E(S|X=1)\times \Pr(X=1) = 0$$

$$E(S^2) = E(S^2|X=0)\times \Pr(\text{X}=0) + E(S^2|X=1)\times \Pr(X=1) = 16\times 0.1 + 4\times 0.9 = 5.2$$

$$E(S^3) = E(S^3|X=0)\times \Pr(\text{X}=0) + E(S^3|X=1)\times \Pr(X=1) = 0$$

$$E(S^4) = E(S^4|X=0)\times \Pr(\text{X}=0) + E(S^4|X=1)\times \Pr(X=1)$$
$$= 3\times 16^2 \times 0.1 + 3\times 4^2 \times 0.9 = 120$$

(e) $\mu_S = E(S) = 0$, thus, $E(S-\mu_S)^3 = E(S^3) = 0$ from part (d). Thus, skewness = 0.

Similarly, $\sigma_S^2 = E(S-\mu_S)^2 = E(S^2) = 5.2$, and $E(S-\mu_S)^4 = E(S^4) = 120$. Thus, $\text{kurtosis} = 120/(5.2^2) = 4.44$.

2.14. The central limit theorem suggests that when the sample size (n) is large, the distribution of the sample average $(\overline{Y})$ is approximately $N\left(\mu_Y, \sigma_{\overline{Y}}^2\right)$ with $\sigma_{\overline{Y}}^2 = \frac{\sigma_Y^2}{n}$. Given $\mu_Y = 50$, $\sigma_Y^2 = 21$,

(a) $n = 50$ and $\sigma_{\overline{Y}}^2 = \frac{\sigma_Y^2}{n} = \frac{21}{50} = 0.42$

$$\Pr(\overline{Y} \leq 51) = \Pr\left(\frac{Y - 50}{\sqrt{0.42}} \leq \frac{51 - 50}{\sqrt{0.42}}\right) = \Pr\left(\frac{Y - 50}{\sqrt{0.42}} \leq 1.5429\right) \approx \Phi(1.543) = 0.9382$$

(b) $n = 150$ and $\sigma_{\overline{Y}}^2 = \frac{\sigma_Y^2}{n} = \frac{21}{150} = 0.14$

$$\begin{aligned}
\Pr(\overline{Y} > 49) &= 1 - \Pr(\overline{Y} \leq 49 \\
&= 1 - \Pr\left(\frac{\overline{Y} - 50}{\sqrt{0.14}} \leq \frac{49 - 50}{\sqrt{0.14}}\right) \\
&= 1 - \Pr\left(\frac{\overline{Y} - 50}{\sqrt{0.14}} \geq \frac{-1}{0.3741}\right) \\
&= 1 - \Pr\left(\frac{\overline{Y} - 50}{\sqrt{0.14}} \geq -2.6731\right) \approx 1 - \Phi(-2.6731) \\
&= \Phi(2.6731) = 0.9962
\end{aligned}$$

(c) $n = 45$ and $\sigma_{\overline{Y}}^2 = \frac{\sigma_Y^2}{n} = \frac{21}{45} = 0.467$

$$\begin{aligned}
&\Pr(50.5 \leq \overline{Y} \leq 51) \\
&= \Pr\left(\frac{50.5 - 50}{\sqrt{0.467}} \leq \frac{\overline{Y} - 50}{\sqrt{0.467}} \leq \frac{51 - 50}{\sqrt{0.467}}\right) \\
&\approx \Phi(1.4587) - \Phi(0.7293) = 0.9265 - 0.7642 \\
&= 0.1623
\end{aligned}$$

2.15. (a)

$$\Pr(9.6 \le \overline{Y} \le 10.4) = \Pr\left(\frac{9.6-10}{\sqrt{4/n}} \le \frac{\overline{Y}-10}{\sqrt{4/n}} \le \frac{10.4-10}{\sqrt{4/n}}\right)$$
$$= \Pr\left(\frac{9.6-10}{\sqrt{4/n}} \le Z \le \frac{10.4-10}{\sqrt{4/n}}\right)$$

where $Z \sim N(0, 1)$. Thus,

(i) $n = 20$; $\Pr\left(\frac{9.6-10}{\sqrt{4/n}} \le Z \le \frac{10.4-10}{\sqrt{4/n}}\right) = \Pr(-0.89 \le Z \le 0.89) = 0.63$

(ii) $n = 100$; $\Pr\left(\frac{9.6-10}{\sqrt{4/n}} \le Z \le \frac{10.4-10}{\sqrt{4/n}}\right) = \Pr(-2.00 \le Z \le 2.00) = 0.954$

(iii)$n = 1000$; $\Pr\left(\frac{9.6-10}{\sqrt{4/n}} \le Z \le \frac{10.4-10}{\sqrt{4/n}}\right) = \Pr(-6.32 \le Z \le 6.32) = 1.000$

(b)

$$\Pr(10 - c \le \overline{Y} \le 10 + c) = \Pr\left(\frac{-c}{\sqrt{4/n}} \le \frac{\overline{Y}-10}{\sqrt{4/n}} \le \frac{c}{\sqrt{4/n}}\right)$$
$$= \Pr\left(\frac{-c}{\sqrt{4/n}} \le Z \le \frac{c}{\sqrt{4/n}}\right).$$

As n get large $\frac{c}{\sqrt{4/n}}$ gets large, and the probability converges to 1.

(c) This follows from (b) and the definition of convergence in probability given in Key Concept 2.6.

2.16. There are several ways to do this. Here is one way. Generate n draws of Y, Y_1, Y_2, … Y_n. Let $X_i = 1$ if $Y_i < 3.6$, otherwise set $X_i = 0$. Notice that X_i is a Bernoulli random variables with $\mu_X = \Pr(X = 1) = \Pr(Y < 3.6)$. Compute $\bar{X}$. Because $\bar{X}$ converges in probability to $\mu_X = \Pr(X = 1) = \Pr(Y < 3.6)$, $\bar{X}$ will be an accurate approximation if n is large.

2.17. $\mu_Y = 0.6$ and $\sigma_Y^2 = 0.4 \times 0.6 = 0.24$

(a) (i) $P(\overline{Y} \geq 0.64) =$

$$1 - \Pr\left(\frac{\overline{Y} - 0.6}{\sqrt{0.24/n}} \geq \frac{0.64 - 0.6}{\sqrt{0.24/n}}\right) = 1 - \Pr\left(\frac{\overline{Y} - 0.6}{\sqrt{0.24/n}} \leq 0.5773\right) = 0.72$$

(ii) $P(\overline{Y} \leq 0.56) = \Pr\left(\frac{\overline{Y} - 0.6}{\sqrt{0.24/n}} \leq \frac{0.56 - 0.6}{\sqrt{0.24/n}}\right) = \Pr\left(\frac{\overline{Y} - 0.6}{\sqrt{0.24/n}} \leq -1.154\right) = 0.12$

b) We know Pr(−1.96 ≤ Z ≤ 1.96) = 0.95; thus, we want n to satisfy

$$0.61 = \frac{0.65 - 0.60}{\sqrt{0.24/n}} > -1.96 \text{ and } \frac{0.56 - 0.60}{\sqrt{0.24/n}} < -1.96$$

Solving these inequalities yields $n \geq 368$.

2.18. $\Pr(Y = \$0) = 0.95,\ \Pr(Y = \$30{,}000) = 0.05$

(a) The mean of Y is

$$\mu_Y = 0 \times \Pr(Y = \$0) + 30{,}000 \times \Pr(Y = 530{,}00) = \$1{,}500$$

The variance of Y is

$$\begin{aligned}\sigma_Y^2 &= E\left[\left(Y - \mu_Y\right)^2\right] \\ &= (0-1500)^2 \times \Pr\left(Y = 50\right) + (30000-1500)^2 \times \Pr(Y = \$30000) \\ &= (-1500)^2 \times 0.95 + (28{,}500)^2 \times 0.05 \\ &= 4.27 \times 10^7\end{aligned}$$

so the standard deviation of Y is $\sigma_Y = (4.27 \times 10^7)^{\frac{1}{2}} = \$6{,}538.35$

(b) $\sigma_{\bar{Y}}^2 = \frac{\sigma_Y^2}{n} = 4.27 \times \dfrac{10^7}{120} = 355{,}833$

Using the central limit theorem,

$$\begin{aligned}\Pr(\bar{Y} > 3000) &= 1 - \Pr(\bar{Y} \le 3000) \\ &= 1 - \Pr\left(\frac{Y - 1500}{\sqrt{355{,}833}} \le \frac{3000-1500}{\sqrt{355{,}833}}\right) \\ &\approx 1 - \Phi(2.5145) = 1 - 0.9939 = 0\end{aligned}$$

2.19. (a)

$$\begin{aligned}\Pr(Y=y_j) &= \sum_{i=1}^{l}\Pr(X=x_i, Y=y_j)\\ &= \sum_{i=1}^{l}\Pr(Y=y_j|X=x_i)\Pr(X=x_i)\end{aligned}$$

(b)

$$\begin{aligned}E(Y) &= \sum_{j=1}^{k} y_j \Pr(Y=y_j) = \sum_{j=1}^{k} y_j \sum_{i=1}^{l}\Pr(Y=y_j|X=x_i)\Pr(X=x_i)\\ &= \sum_{i=1}^{l}\left(\sum_{j=1}^{k} y_j \Pr(Y=y_j|X=x_i)\right)\Pr(X=x_i)\\ &= \sum_{i=1}^{l} E(Y|X=x_i)\Pr(X=x_i).\end{aligned}$$

(c) When X and Y are independent,

$$\Pr(X=x_i, Y=y_j) = \Pr(X=x_i)\Pr(Y=y_j),$$

so

$$\begin{aligned}\sigma_{XY} &= E[(X-\mu_X)(Y-\mu_Y)]\\ &= \sum_{i=1}^{l}\sum_{j=1}^{k}(x_i-\mu_X)(y_j-\mu_Y)\Pr(X=x_i, Y=y_j)\\ &= \sum_{i=1}^{l}\sum_{j=1}^{k}(x_i-\mu_X)(y_j-\mu_Y)\Pr(X=x_i)\Pr(Y=y_j)\\ &= \left(\sum_{i=1}^{l}(x_i-\mu_X)\Pr(X=x_i)\right)\left(\sum_{j=1}^{k}(y_j-\mu_Y)\Pr(Y=y_j\right)\\ &= E(X-\mu_X)E(Y-\mu_Y) = 0\times 0 = 0,\end{aligned}$$

$$cor(X,Y) = \frac{\sigma_{XY}}{\sigma_X\sigma_Y} = \frac{0}{\sigma_X\sigma_Y} = 0.$$

2.20. (a) $\Pr(Y = y_i) = \sum_{j=1}^{l}\sum_{h=1}^{m} \Pr(Y = y_i | X = x_j,\, Z = z_h)\Pr(X = x_j,\, Z = z_h)$

(b)

$$\begin{aligned} E(Y) &= \sum_{i=1}^{k} y_i \Pr(Y = y_i)\Pr(Y = y_i) \\ &= \sum_{i=1}^{k} y_i \sum_{j=1}^{l}\sum_{h=1}^{m} \Pr(Y = y_i | X = x_j,\, Z = z_h)\Pr(X = x_j,\, Z = z_h) \\ &= \sum_{j=1}^{l}\sum_{h=1}^{m}\left[\sum_{i=1}^{k} y_i \Pr(Y = y_i | X = x_j,\, Z = z_h)\right]\Pr(X = x_j,\, Z = z_h) \\ &= \sum_{j=1}^{l}\sum_{h=1}^{m} E(Y | X = x_j,\, Z = z_h)\Pr(X = x_j,\, Z = z_h) \end{aligned}$$

where the first line in the definition of the mean, the second uses (a), the third is a rearrangement, and the final line uses the definition of the conditional expectation.

2. 21.

(a)

$$\begin{aligned} E(X-\mu)^3 &= E[(X-\mu)^2(X-\mu)] = E[X^3 - 2X^2\mu + X\mu^2 - X^2\mu + 2X\mu^2 - \mu^3] \\ &= E(X^3) - 3E(X^2)\mu + 3E(X)\mu^2 - \mu^3 \\ &= E(X^3) - 3E(X^2)E(X) + 3E(X)[E(X)]^2 - [E(X)]^3 \\ &= E(X^3) - 3E(X^2)E(X) + 2E(X)^3 \end{aligned}$$

(b)

$$\begin{aligned} E(X-\mu)^4 &= E[(X^3 - 3X^2\mu + 3X\mu^2 - \mu^3)(X-\mu)] \\ &= E[X^4 - 3X^3\mu + 3X^2\mu^2 - X\mu^3 - X^3\mu + 3X^2\mu^2 - 3X\mu^3 + \mu^4] \\ &= E(X^4) - 4E(X^3)E(X) + 6E(X^2)E(X)^2 - 4E(X)E(X)^3 + E(X)^4 \\ &= E(X^4) - 4[E(X)][E(X^3)] + 6[E(X)]^2[E(X^2)] - 3[E(X)]^4 \end{aligned}$$

2. 22. The mean and variance of R are given by

$$\mu = w\times 0.08 + (1-w)\times 0.05$$
$$\sigma^2 = w^2\times 0.07^2 + (1-w)^2\times 0.042 + 2\times w\times (1-w)\times[0.07\times 0.04\times 0.25]$$

where $0.07\times 0.04\times 0.25 = Cov(R_s,\ R_b)$ follows from the definition of the correlation between R_s and R_b.

(a) $\mu = 0.065;\ \sigma = 0.044$

(b) $\mu = 0.0725;\ \sigma = 0.056$

(c) $w = 1$ maximizes μ; $\sigma = 0.07$ for this value of w.

(d) The derivative of σ^2 with respect to w is

$$\frac{d\sigma^2}{dw} = 2w\times .07^2 - 2(1-w)\times 0.04^2 + (2-4w)\times[0.07\times 0.04\times 0.25]$$
$$= 0.0102w - 0.0018$$

Solving for w yields $w = 18/102 = 0.18.$ (Notice that the second derivative is positive, so that this is the global minimum.) With $w = 0.18,\ \sigma_R = .038.$

2. 23. X and Z are two independently distributed standard normal random variables, so

$$\mu_X = \mu_Z = 0, \sigma_X^2 = \sigma_Z^2 = 1, \sigma_{XZ} = 0.$$

(a) Because of the independence between X and Z, $\Pr(Z = z|X = x) = \Pr(Z = z)$, and $E(Z|X) = E(Z) = 0$. Thus
$E(Y|X) = E(X^2 + Z|X) = E(X^2|X) + E(Z|X) = X^2 + 0 = X^2.$

(b) $E(X^2) = \sigma_X^2 + \mu_X^2 = 1,$ and $\mu_Y = E(X^2 + Z) = E(X^2) + \mu_Z = 1 + 0 = 1.$

(c) $E(XY) = E(X^3 + ZX) = E(X^3) + E(ZX).$ Using the fact that the odd moments of a standard normal random variable are all zero, we have $E(X^3) = 0.$ Using the independence between X and Z, we have $E(ZX) = \mu_Z \mu_X = 0.$ Thus
$E(XY) = E(X^3) + E(ZX) = 0.$

(d)

$$\begin{aligned}\operatorname{cov}(XY) &= E[(X - \mu_X)(Y - \mu_Y)] = E[(X - 0)(Y - 1)] \\ &= E(XY - X) = E(XY) - E(X) \\ &= 0 - 0 = 0.\end{aligned}$$

$$\operatorname{corr}(X, Y) = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} = \frac{0}{\sigma_X \sigma_Y} = 0.$$

2.24. (a) $E(Y_i^2) = \sigma^2 + \mu^2 = \sigma^2$ and the result follows directly.

(b) (Y_i/σ) is distributed i.i.d. $N(0,1)$, $W = \sum_{i=1}^{n}(Y_i/\sigma)^2$, and the result follows from the definition of a χ_n^2 random variable.

(c) $E(W) = E\sum_{i=1}^{n}\frac{Y_i^2}{\sigma^2} = \sum_{i=1}^{n}E\frac{Y_i^2}{\sigma^2} = n.$

(d) Write

$$V = \frac{Y_1}{\sqrt{\frac{\sum_{i=2}^{n}Y_i^2}{n-1}}} = \frac{Y_1/\sigma}{\sqrt{\frac{\sum_{i=2}^{n}(Y/\sigma)^2}{n-1}}}$$

which follows from dividing the numerator and denominator by σ. $Y_1/\sigma \sim N(0,1)$, $\sum_{i=2}^{n}(Y_i/\sigma)^2 \sim \chi_{n-1}^2$, and Y_1/σ and $\sum_{i=2}^{n}(Y_i/\sigma)^2$ are independent. The result then follows from the definition of the *t* distribution.

2.25. (a) $\sum_{i=1}^{n} ax_i = (ax_1 + ax_2 + ax_3 + \cdots + ax_n) = a(x_1 + x_2 + x_3 + \cdots + x_n) = a\sum_{i=1}^{n} x_i$

(b)

$$\begin{aligned}\sum_{i=1}^{n}(x_i + y_i) &= (x_1 + y_1 + x_2 + y_2 + \cdots x_n + y_n)\\ &= (x_1 + x_2 + \cdots x_n) + (y_1 + y_2 + \cdots y_n)\\ &= \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} y_i\end{aligned}$$

(c) $\sum_{i=1}^{n} a = (a + a + a + \cdots + a) = na$

(d)

$$\begin{aligned}\sum_{i=1}^{n}(a + bx_i + cy_i)^2 &= \sum_{i=1}^{n}(a^2 + b^2x_i^2 + c^2y_i^2 + 2abx_i + 2acy_i + 2bcx_iy_i)\\ &= na^2 + b^2\sum_{i=1}^{n} x_i^2 + c^2\sum_{i=1}^{n} y_i^2 + 2ab\sum_{i=1}^{n} x_i + 2ac\sum_{i=1}^{n} y_i + 2bc\sum_{i=1}^{n} x_iy_i\end{aligned}$$

2.26. (a) corr(Y_i,Y_j) = $\frac{\text{cov}(Y_i, Y_j)}{\sigma_{Y_i}\sigma_{Y_j}} = \frac{\text{cov}(Y_i, Y_j)}{\sigma_Y \sigma_Y} = \frac{\text{cov}(Y_i, Y_j)}{\sigma_Y^2} = \rho$, where the first equality uses the definition of correlation, the second uses the fact that Y_i and Y_j have the same variance (and standard deviation), the third equality uses the definition of standard deviation, and the fourth uses the correlation given in the problem. Solving for cov(Y_i, Y_j) from the last equality gives the desired result.

(b) $\bar{Y} = \frac{1}{2}Y_1 + \frac{1}{2}Y_2$, so that $E(\bar{Y}) = \frac{1}{2}E(Y)_1 + \frac{1}{2}E(Y_2) = \mu_Y$

$$\text{var}(\bar{Y}) = \frac{1}{4}\text{var}(Y_1) + \frac{1}{4}\text{var}(Y_2) + \frac{2}{4}\text{cov}(Y_1, Y_2) = \frac{\sigma_Y^2}{2} + \frac{\rho\sigma_Y^2}{2}$$

(c) $\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$, so that $E(\bar{Y}) = \frac{1}{n}\sum_{i=1}^{n} E(Y_i) = \frac{1}{n}\sum_{i=1}^{n}\mu_Y = \mu_Y$

$$\begin{aligned}
\text{var}(\bar{Y}) &= \text{var}\left(\frac{1}{n}\sum_{i=1}^{n} Y_i\right) \\
&= \frac{1}{n^2}\sum_{i=1}^{n}\text{var}(Y_i) + \frac{2}{n^2}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\text{cov}(Y_i, Y_j) \\
&= \frac{1}{n^2}\sum_{i=1}^{n}\sigma_Y^2 + \frac{2}{n^2}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\rho\sigma_Y^2 \\
&= \frac{\sigma_Y^2}{n} + \frac{n(n-1)}{n^2}\rho\sigma_Y^2 \\
&= \frac{\sigma_Y^2}{n} + \left(1 - \frac{1}{n}\right)\rho\sigma_Y^2
\end{aligned}$$

where the fourth line uses $\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} a = a(1+2+3+\cdots+n-1) = \frac{an(n-1)}{2}$ for any variable a.

(d) When n is large $\frac{\sigma_Y^2}{n} \approx 0$ and $\frac{1}{n} \approx 0$, and the result follows from (c).

2.27

(a) $E(W) = E[E(W|Z)] = E[E(X - \tilde{X})|Z] = E[E(X|Z) - E(X|Z)] = 0$.

(b) $E(WZ) = E[E(WZ|Z)] = E[ZE(W)|Z] = E[Z \times 0] = 0$

(c) Using the hint: $V = W - h(Z)$, so that $E(V^2) = E(W^2) + E[h(Z)^2] - 2 \times E[W \times h(Z)]$. Using an argument like that in (b), $E[W \times h(Z)] = 0$. Thus, $E(V^2) = E(W^2) + E[h(Z)^2]$, and the result follows by recognizing that $E[h(Z)^2] \geq 0$ because $h(z)^2 \geq 0$ for any value of z.